



CAPSTONE PROJECT

Mentor : Dr. Ram Narayan Yadav

REAL TIME ANALYSIS OF STREAMING DATA USING APACHE STORM



Group members:

Prashanth Kurella
Ojasvi Bhalerao
Abhinav Ralhan
Gaurav Malhotra
Anirudh Kumar Rao

Objectives

Objective :

- Live analysis on twitter stream data and connect it to a dashboard.

Background

The importance of this project is to understand the working of real-time data analysis tools like Apache Storm.

What is the use of analytics ?

Storm vs Hadoop ?

Why use Storm ?

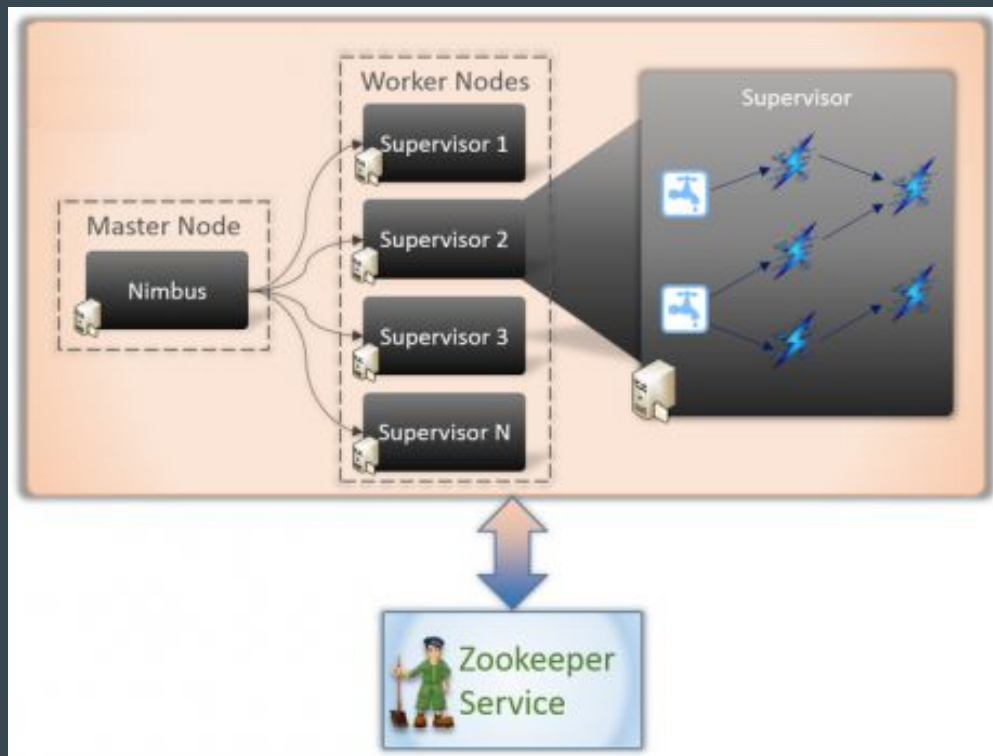
Apache Storm

Apache storm is an open-source distributed real-time computational system for processing data streams. Similar to what Hadoop used to do for batch processing, Apache Storm does for unbounded streams of data in a reliable manner.

- Able to process over a million jobs in fraction of a second on a node
- Integrated with Hadoop to harness higher throughputs
- Easy to implement and can be integrated with any programming language



Apache Storm



Thank you